# Good Firms, Worker Flows and Productivity *

Michel Serafinelli†

University of California, Berkeley

JOB MARKET PAPER

November 2012

**Abstract**

A consensus has emerged that agglomeration economies can at least partially explain why firms cluster next to each other. Disagreement remains, however, over the sources of these agglomeration effects. In this paper I present direct empirical evidence on the role of firm-to-firm labor mobility in enhancing the productivity of firms located near other highly productive firms. Using matched employer-employee and balance sheet data for Veneto, a region of Italy with many successful industry clusters, I first identify a set of high-wage firms (HWF). I show that these firms have higher labor productivity and higher intangible capital per worker than other firms in the same industry. I then show that hiring a worker with HWF experience increases the productivity of other (non-HWF) firms by 3-5 percent on average. This productivity effect is not driven by shocks that lead to the hiring workers in general, nor is it attributable to unobserved productivity shocks that are correlated with the propensity to hire workers with HWF experience. Back-of-the-envelope calculations suggest that worker flows can explain 15-20 percent of the productivity gains experienced by other firms when high-productivity plants are added to a local labor market.

*I am very grateful to David Card, Patrick Kline and Enrico Moretti for invaluable guidance and support throughout the duration of this project, and to Bronwyn Hall for very helpful suggestions. I also thank Andrés Rodríguez-Clare, Yuriy Gorodnichenko, Miguel Almunia, Vladimir Asriyan, Audinga Baltrunaite, Pamela Campa, Frederico Finan, Francesco Devicienti, Tadeja Gracner, Jonas Hjort, Agata Maida, Steven Raphael, Ana Rocca, Jesse Rothstein, Fabio Schiantarelli, Victoria Vanasco, and seminar participants at Berkeley Econ, Berkeley Haas, Milan MILLS, Stockholm University, EIEF Rome MOOD, San Francisco WEAI, Toronto RCEF, All-California Labor Conference at RAND and San Francisco Fed for comments. I acknowledge financial support from the Center for Equitable Growth at UC Berkeley. Thanks are due also to Veneto firms' employees and officials of employers' associations and chambers of commerce (in particular Federico Callegari) who have given generously their time in interviews. I am indebted to Giuseppe Tattara for making available the Veneto Work History Data. I am grateful to Bocconi University library staff for assistance with the AIDA data and to Harrison Dekker and John Ridener for programming help. This research also benefited from the support and services of UC Berkeley's Geospatial Innovation Facility (GIF). Nitin Kohli provided excellent research assistance. The latest version of this paper is available at http://econgrads.berkeley.edu/serafine/.

†Email: serafine@econ.berkeley.edu

# 1 Introduction

A prominent feature of the economic landscape in the most developed countries is the tendency for firms to locate near other firms producing similar products or services. In the United States, for example, biopharmaceutical firms are clustered in New York and Chicago, the carpet and rug industry is concentrated in the area around Dalton, Georgia, and a sizeable share of the aircraft engine industry is clustered around Hartford, Connecticut. In addition, the growth and diffusion of multinational corporations has led to the recent appearance of important industrial clusters in several emerging economies. Firms that originally agglomerated in Silicon Valley and Detroit now have subsidiaries clustered in Bangalore and Slovakia (Alfaro and Chen, 2010).

Researchers have long speculated that firms in industrial concentrations may benefit from agglomeration economies, and a growing body of work has been devoted to studying the importance of these economies. Despite the difficulties involved in estimating agglomeration effects, a consensus has emerged from the literature that significant productivity advantages of agglomeration exist for many industries (Henderson, 2003; Rosenthal and Strange, 2003; Ellison, Glaeser and Kerr, 2010; Cingano and Schivardi, 2004; Overman and Puga, 2010; Greenstone, Hornbeck and Moretti, 2010). Disagreement remains, however, over the identity and relative importance of alternative mechanisms that can account for these advantages (Glaeser and Gottlieb, 2009; Moretti, 2011). Knowledge spillovers, labor market pooling and the availability of specialized intermediate inputs have garnered the most attention in the literature's attempt to explain agglomeration economies.[1]

This paper directly examines the role of labor mobility as a mechanism for the transfer of efficiency-enhancing knowledge and evaluates the extent to which labor mobility can explain the existing evidence on the productivity advantages of agglomeration. Instead of considering knowledge spillovers and labor market pooling as distinct source of agglomeration advantages, as has been done in the empirical literature to date, I consider them jointly. The underlying idea is that knowledge is embedded in workers and diffuses when workers move between firms. The strong localized aspect of knowledge spillovers discussed in the agglomeration literature may thus arise from the propensity of workers to change jobs within the same local labor market.[2]

[1]For theoretical analysis of knowledge spillovers, see Fujita and Ogawa (1982), Helsley (1990) and Glaeser (1999). For analysis of labor market pooling, see Kim (1999), Helsley and Strange (1990), Rotenberg and Saloner (2000) and Acemoglu (1997). For analysis of the availability of specialized intermediate inputs, see Abdel, Rahman and Fujita (1990). Duranton and Puga (2004) surveys the theoretical literature on micro-foundations for agglomeration advantages.

[2]Recently, Combes and Duranton (2006) have reconsidered knowledge spillovers and labor market pooling from a theoretical standpoint. One of the main implications of their model is that spillovers

Identifying the microeconomic mechanisms underlying localized productivity spillovers is crucial for understanding agglomeration economies. Without knowing the precise nature of the interactions between firms and workers that generate agglomeration advantages, it is difficult to be confident about the existence of any such advantages. Additionally, pinpointing the ultimate causes of agglomeration advantages is helpful for understanding differences in productivity across industry clusters and localities. Finally, better knowledge of the sources of the productivity advantages of agglomeration is important for determining the optimal design of location-based policies (Kline, 2010; Kline and Moretti, 2011).

In order to empirically assess the importance of labor-market based knowledge spillovers, I use a unique dataset from the Veneto region of Italy that combines Social Security earnings records and detailed financial information for firms. While the issues analyzed in this paper are of general interest, the case of Veneto is important because this region is part of a larger economic area of Italy where, as in the Silicon Valley, networks of specialized firms have been effective in promoting and adapting to technological change during the last three decades. This so called "Third Italy" region has received a good deal of attention by researchers, in the United States as well as in Europe (Piore and Sabel, 1984; Piore, 2009; Brusco, 1982; Trigilia, 1990; Whitford, 2001).

I begin by presenting a simple conceptual framework where some firms are more productive because they have some superior knowledge.[3] Employees at these firms passively acquire some proportion of the firm's internal knowledge. For simplicity, I refer to these as "knowledgeable" workers. Other firms can gain access to the superior knowledge by hiring these workers. Empirically, I identify potentially high-productivity firms as those that pay a relatively high firm-specific wage premium.[4]

I show that these high-wage-firms (HWFs) have higher labor productivity, higher total factor productivity, and higher capital (in particular intangible capital) per worker, suggesting the presence of a firm-specific productivity advantage and thus a point of origin for the transfer of knowledge. Next, I evaluate the extent to which non-HWFs benefit from hiring knowledgeable workers by studying the effect on productivity associated with hiring workers with recent experience at HWFs.

An obvious concern is that firms that hire workers with recent HWF experience are dif-

---

and labor market pooling should not be viewed as separate mechanisms for productivity advantages of agglomeration, since the local labor market may function as a conduit for the diffusion of information.

[3] I interpret knowledge broadly in this paper and allow its definition to include information about export markets, physical capital, process innovations, new managerial techniques, new organizational forms and intermediate inputs.

[4] This is consistent with many recent models of frictional labor markets (e.g., Christensen et al., 2005), in which higher-productivity firms pay higher wages for equivalent workers

ferent than those that do not, and that this underlying difference – rather than knowledge transfers – account for the measured productivity effects. My procedure for measuring the number of workers with recent HWF experience focuses on newly hired workers. If workers who change establishments are more productive than stayers in general, or if the number of recent hires is systematically correlated with time-varying unobservables at the firm level, OLS estimates of the effect of hiring workers with recent HWF experience will overestimate the importance of labor mobility for knowledge transfer. In order to address this issue, I augment the regression equation of interest with the number of recently hired workers without experience at HWFs. By including both types of new hires, any potential bias caused by the correlation between time-varying unobserved productivity shocks and hiring in general is removed.

Productivity shocks that are correlated with the propensity to hire knowledgeable workers may also give rise to an upward bias in the differential impact of knowledgeable workers. This problem is known as 'transmission bias,' since these shocks are transmitted to specific input choices (Eberhardt and Helmer, 2010). In order to address this potential endogeneity issue, I use standard control function methods from the recent productivity literature (Olley and Pakes, 1996; Levinsohn and Petrin, 2003). I conclude that the average effect of recruiting a knowledgeable worker on a non-HWF's productivity is an increase of between 3 and 5 percent. This productivity effect of knowledgeable workers is not associated with recently hired workers in general; I do not find a similar productivity effect for recently hired workers without experience at good firms.

Another potential threat to identification is the fact that I do not observe labor quality. Of particular concern are workers who might separate from an HWF because their level of ability is low. I refer to this potential adverse selection problem as "lemons bias" (Gibbons and Katz, 1991). Lemons bias will tend to work against the finding of a positive effect of knowledgeable workers. In order to address this issue, I obtain a proxy for worker ability and I weight the number of workers in my OLS regression using the average ability to obtain effective labor input.[5] To further guard against the possibility of lemons bias, I instrument for the number of knowledgeable workers in a non-HWF with the number of local good firms in the same industry that downsized in the previous period. When a larger number of workers are being laid off from HWFs the extent of lemons' bias is arguably reduced. The IV estimates return an economically and statistically significant effect of recruiting knowledgeable workers on non-HWF productivity, with the point estimate larger than the OLS. While in principle this is consistent with the idea that the OLS

[5] To obtain this proxy for ability, I procure estimates of worker fixed effects from wage equations where both firm and worker effects can be identified.

coefficient is biased downward, in practice the IV standard errors are large and prevent me from drawing definitive conclusions.[6]

In the last part of the paper, I assess the extent to which worker flows can explain the productivity advantages of agglomeration. I relate my findings to the existing evidence on the productivity advantages of agglomeration, focusing in particular on the study performed in Greenstone, Hornbeck and Moretti (2010, henceforth GHM). GHM find that after the opening of a large manufacturing establishment, total factor productivity (TFP) of incumbent plants in US counties that were able to attract one of these large plants increases significantly relative to the TFP of incumbent plants in counties that survived a long selection process but narrowly lost the competition. Moreover, the observed effect on TFP is larger if incumbent plants are in the same industry as the large plant, and increases over time. These two facts are consistent with the presence of intellectual externalities that are embodied in workers who move from firm to firm. While their research design allows GHM to obtain credible estimates of the effect of the large plant's entry on local TFP, data limitations prevent them from drawing definitive conclusions regarding the driving mechanism. I evaluate the extent to which worker flows explain empirical evidence on the productivity advantages of agglomeration, by simulating an event similar to that studied by GHM but within the worker mobility framework described above. The change in productivity predicted within this framework equals 15-20 percent of the overall effect found in GHM, suggesting that knowledge transfer through worker flows explain an important portion of agglomeration advantages.

The remainder of this paper is structured as follows. In Section 2, I relate my research to the existing literature. Section 3 presents a production function framework that guides the empirical exercise and aids in interpreting the results. Section 4 presents the econometric model and discusses relevant estimation issues. In Section 5, I describe my data and provide a descriptive overview. The regression results, in addition to various extensions and robustness checks are presented in Section 6 and 7. Section 8 concludes the paper.

# 2 Relation to Previous Research

This paper adds to the growing literature on productivity advantages through agglomeration, a literature critically surveyed in Rosenthal and Strange (2004), Duranton and Puga

[6] Another tentative explanation for the magnitude of the IV results is that recruiting a knowledgeable worker may have higher productivity effect for the subgroup of firms that are marginal in the sense that they recruit knowledgeable workers if and only if there exists excess local supply.

(2004) and Moretti (2011). The research relating most closely to this paper is the body of work on micro-foundations for agglomeration advantages based on knowledge spillovers and labor market pooling. Localized knowledge spillovers are a common explanation for the productivity advantages of agglomeration. Nevertheless, as pointed out by Combes and Duranton (2006), if information can easily flow out of firms, it must be clarified why the effects of spillovers are localized. Duranton and Puga (2004) present a model, inspired by Jovanovic and Rob (1989), Jovanovic and Nyarko (1995) and Glaeser (1999), in which proximity to individuals with greater skills or knowledge facilitates the acquisition of skills and the exchange and diffusion of knowledge. Building on the work of Fujita and Ogawa (1982), Helsley (1990) proposes a model where the knowledge produced in a location is a by-product of output, and diffuses through contacts between firms whose cost rises with distance. However, it remains unclear what are the frictions associated with the spatial propagation of information.

For labor market pooling, the argument is that agglomeration allows a better match between employer needs and worker skills, which may result in higher productivity (Kim, 1989, Helsley and Strange, 1990). Rotenberg and Saloner (2000) argue that, by hosting a large number of potential partners, large cities or industrial concentrations can help mitigate hold-up problems that plague bilateral relationships between employers and employees. Specifically, the competition between firms to hire skilled workers makes it easier for skilled workers to recoup the cost of acquiring industry-specific human capital. Moreover, Acemoglu (1997) maintains that in large cities or industrial concentrations, firms invest in new technologies because they know that they can find specialized employees.[7] While labor market pooling is a potentially promising explanation for the productivity advantages of agglomeration, the existing evidence is still limited and rather indirect.

The theoretical analysis in Combes and Duranton (2006) also explores issues related to the empirical exercise in this paper. More precisely, the authors argue that firms clustering in the same locality face a trade-off between the advantages of labor pooling (i.e. access to knowledge carriers) and the costs of labor poaching (i.e. loss of some key employees to competitors along with higher wage bills to retain other key employees). In the context of a duopoly game, the authors illustrate how firms' strategic decisions regarding locations, wages, poaching and prices depend on market size, the degree of horizontal differentiation between goods, and on worker heterogeneity in terms of knowledge transfer cost.[8] In

[7] The model outlined by Kim (1989) also predicts that workers in a larger market invest more in the depth of their human capital and less in the width.

[8] Additionally, the study of R&D spillover effects by Bloom, Schankerman, and Van Reenen (2007) points out the presence of two countervailing effects: positive technological spillovers and negative business-stealing effects on the product market. The authors provide evidence that altough both types

her influential book, Saxenian (1994) argue that the geographic proximity of high-tech firms in Silicon Valley is associated with a more efficient flow of new ideas. Specifically, Saxenian argues [p. 37] that

> The decentralized and fluid environment accelerated the diffusion of technological capabilities and know-how within the region... When engineers moved between companies, they took with them the knowledge, skills, and experience acquired at their previous jobs.

I contribute to the literature on micro-foundations for agglomeration advantages by showing direct evidence of knowledge transfer using a "paper trail" based on the movement of workers between firms in my matched data to test the extent to which knowledge spillovers are geographically localized. I find support for agglomeration advantages through worker flows. My results may also help explain the findings in Henderson (2003), Cingano and Schivardi (2004) and Moretti (2004b) that local spillovers are increasing in economic proximity.

Some research beyond the agglomeration literature has also emphasized the fact that new workers share ideas on how to organize production or information on new technologies that they learned with their previous employer. Theoretical studies on this theme include Fosfuri, Motta and Rønde (2001), Cooper (2001), Markusen (2001), Glass and Saggi (2002), Gersbach and Schmutzler (2003) and Dasgupta (2012). Empirical work includes that of Song, Almeida and Wu (2003), which shows that labor turnover can explain patterns of patent citations. Other empirical studies, including Rao and Drazin (2002), Kaiser, Kongsted and Rønde (2008) and Maliranta, Mohnen and Rouvinen (2009) find that firms hiring workers from R&D-intensive firms tend to perform better. Görg and Strobl (2005) show that domestic firms established in Ghana by entrepreneurs with experience from foreign-owned companies in the same industry are more productive and more likely to survive than firms established by entrepreneurs with no experience at foreign-owned companies. In general, research on spillovers from foreign to domestic firms has expanded the scope of knowledge spillovers by looking at a broader knowledge than that possessed and transferred by R&D labor alone.[9] Using plant-level data from Colombia, Markusen and Trofimenko (2009) present evidence to support the hypothesis that the use of foreign experts have substantial, although not always immediate, positive effects on the value added per worker. Balsvik (2011) uses matched employer-employee data from

of effects operate, technological spillovers quantitatively dominate.

[9] The discussion in this section centers on research involving labor mobility. For a comprehensive survey of the large empirical literature on spillovers from foreign direct investment to host country firms, see Görg and Greenaway (2004)

Norway and offers a detailed account of productivity gains linked to worker flows from foreign multinational to domestic firms.[10] In a similar vein, using linked worker-firm data, Parrotta and Pozzoli (2012) and Stoyanov and Zubanov (2012) show evidence from Denmark that is consistent with models of knowledge diffusion through labor mobility.[11] The findings in my paper are similar to those of Balsvik (2011), Parrotta and Pozzoli (2012) and Stoyanov and Zubanov (2012). My empirical strategy, however, allows me to more convincingly identify the causal effect of recruiting knowledgeable workers on productivity. To the best of my knowledge, this paper is the first to address both transmission bias and the lemons bias using a combination of productivity literature techniques and an IV approach. Furthermore, while the above authors focus exclusively on the role of labor mobility for knowledge transfer, I seek to shed light on a broader question: the extent to which labor mobility can explain evidence on the productivity advantages through agglomeration.

# 3 The Production Function Framework

This paper seeks to evaluate the extent to which non-HWFs benefit from hiring workers from HWFs. Thus, I begin by presenting a production function framework that guides the subsequent empirical work and aids in interpreting the results. Assume there exists a finite number of different locations, each constituting a separate local labor market. To fix ideas, assume that these labor markets are completely segmented with workers being immobile between them. There exists a finite collection $\mathcal{J} = \{\mathcal{J}_0, \mathcal{J}_1\}$ of firms consisting of the set $\mathcal{J}_1$ of *good* firms, which are more productive because they have some superior knowledge, and set $\mathcal{J}_0$ of other firms which have no access to the superior knowledge. The superior knowledge is exogenously given and could include information about export markets, physical capital, process innovations, new managerial techniques, new organizational forms and intermediate inputs. Below, I will define the good firms as high-wage-firms (HWFs) and show that they are more productive and have higher levels of intangible capital per worker, suggesting the presence of a firm-specific productivity advantage that could generate a point of origin for knowledge transfer.

Workers employed by good firms acquire some proportion of the firm's internal knowl-

[10] Poole (2009) finds a positive effect on wages paid in domestic firms in Brazil of the share of new workers previously employed by foreign-owned firms.

[11] If non-HWFs can gain access to superior knowledge by hiring knowledgeable workers, one would expect a more concentrated productivity distribution in industries with higher rates of worker turnover from more to less productive firms. Stoyanov and Zubanov (2012) report a strong negative correlation between labor flows from more to less productive firms and productivity variance in 21 two-digit industries.

edge. For simplicity, I assume that this acquisition of internal knowledge takes place immediately after the workers join the good firm. Workers are *knowledgeable* if they have knowledge of the relevant information and *unknowledgeable* otherwise. All workers employed by good firms, then, are knowledgeable. Additionally, if a worker moves to a non-good firm, some proportion of this knowledge can be *transferred* to a $j \in \mathcal{J}_0$ firm if the workers switch employers. [12]Output is a nationally traded good and its price is fixed and normalized to one. Capital is supplied at a fixed rental rate $\rho$ to all localities and industries. The production function of firm $j \in \mathcal{J}_0$ is

$$Y_j = F(L_j, K_j, M_j) = A_j[(\overline{\theta}_j L_j)^{\alpha} K_j^{\gamma} M_j^{\lambda}]^{\delta} \tag{1}$$

where $L = H + N$, i.e. the sum of knowledgeable workers ($H$,who moved at some point from a good firm to a non-good firm) and unknowledgeable workers ($N$); $\overline{\theta}$ is the quality of the workforce, $K$ is total capital inputs, $M$ is material inputs, and $\delta < 1$ represent an element of diminishing return to scale, or to "span of control" in the managerial technology (Lucas, 1978). I allow for knowledge transfer by letting firms' productivity depend on the number of knowledgeable workers in the firm:

$$A_j = D_j e^{\beta_H H_j} \tag{2}$$

# 4 Identification

This section builds the main regression equation from the production function framework and discusses threats to identification. The section concludes with a discussion of estimation methods used to identify the effect of knowledgeable workers on productivity. By combining equation (1) and (2) , and taking logs, I obtain the regression equation that forms the basis of my empirical analysis:

$$\ln(Y_{jst}) = \beta_L \ln(\overline{\theta}_{jt} L_{jst}) + \beta_K \ln(K_{jst}) + \beta_M \ln(M_{jst}) + \beta_H H_{jst} + \beta_0 + \zeta_{jst} \tag{3}$$

where the real value of total firm production $Y$ is the dependent variable, $s$ denotes industry, and $t$ denotes year.[13] The term $\ln(D_j)$ is decomposed into two elements, $\beta_0$ and $\zeta_{jst}$. The constant $\beta_o$ denotes mean efficiency across all firms in $\mathcal{J}_0$ that is due do factors

[12] I assume that this type of knowledge cannot all be patented and that exclusive labor contracts are not available.

[13] $\beta_L = \delta\alpha, \beta_K = \delta\gamma, \beta_M = \delta\lambda$

others than $H$. The time-variant $\zeta_{jst}$ represents deviations from this mean efficiency level and captures (i) unobserved factors affecting firm output, (ii) measurement error in inputs and output, and (iii) random noise. I discuss the nature of $\zeta_{jst}$ in greater detail in the next section.

$H$ is constructed from head counts in the matched employer-employee data. I only include workers with recent experience at good firms in $H$. I define a worker as having recent HWF-experience in year $t$, if he or she is observed working in a HWF for one or more of the years $t-k$ to $t-1$, where $k$ is chosen to equal 8 in the baseline specification.[14] In equation (3) $\beta_H$ is the parameter of interest. It tests for an increase in productivity following an increase in the number of knowledgeable workers. Estimating the effect of recruiting a worker with recent HWF experience on a non-HWF's productivity is difficult in the presence of unobservables correlated with new hires, unobserved idiosyncratic shocks to the receiving firm's productivity, and unobserved labor quality. I turn now to describing what type of biases these unobservables may introduce and how I deal with them in the empirical work.

## 4.1 Unobservables correlated with new hires

The way I have constructed the measure for the $H$ implies that this measure captures the newly hired knowledgeable workers, where "newly hired" means hired in year $t, t-1, ..., t-k$. If workers who change establishments are more productive than stayers in general, the effect of newly hired workers with HWF experience may equally apply to newly hired employees without HWF experience. The OLS also overestimates the importance of labor mobility for knowledge transfer if the number of recent hires is systematically correlated with time-varying unobservables at the firm level.

In order to address this issue, I augment equation (3) with $\widetilde{N}$, the number of recently hired workers without experience from good firms. In this augmented specification, the identification of knowledge transfer relies on the differential effect of hiring an employee with HWF experience over hiring an employee from another non-HWF.[15] By including both $H$ and $\widetilde{N}$, any potential bias caused by the correlation between time-varying unobserved productivity shocks and hiring in general is removed. However, productivity shocks that are correlated with the propensity to hire knowledgeable workers may give rise to an

[14] Choosing k=8 allows me to consider as many events of mobility out of HWFs as possible, given that my dataset begins in 1987 and I run production function regressions from 1995 onward (see Section 5). I experimented with $k = 5, 6, 7$: the results remained largely unchanged.

[15] Balsvik (2011) uses a similar approach by dividing workers newly hired by Norwegian firms into two groups: those with experience from multinational enterprises, and those without any such experience.

upward bias in the differential impact of $H$. This problem is known as 'transmission bias' since the shock is suggested to 'transmit to' input choices (Eberhardt and Helmer, 2010).

## 4.2 Transmission Bias

Express $\zeta_{jst}$ , the deviations from mean firm efficiency not resulting from knowledge transfer , as

$$\zeta_{jst} = \omega^*_{jst} + \nu_{jst} = \omega_{jst} + \mu_{st} + \nu_{jst} \tag{4}$$

which specifies that $\zeta_{jst}$ contains measurement error $\nu_{jst}$ and a productivity component $\omega^*_{jst}$ (TFP) known to the firm but unobserved by the econometrician. The productivity component can be further divided into a a firm-specific term and a term common to all firms in a given industry, where the latter represents shocks affecting all firms in the industry in the same way. [16]I use $\mu_{st}$ to represent these shocks. Equation (3) now becomes:

$$\ln(Y_{jst}) = \beta_0 + \beta_L \ln(\overline{\theta} L_{jst}) + \beta_K \ln(K_{jst}) + \beta_M \ln(M_{jst}) + \beta_H H_{jst} + \mu_{st} + \omega_{jst} + \upsilon_{jst} \tag{5}$$

The main difficulty in estimating $\beta_H$ in Equation (5) is that non-HWFs may decide on their choice of $H$ based on the realized firm-specific productivity shock ($\omega_{jst}$) unknown to the researcher. When employing OLS to estimate Equation (5) without accounting for the existence of $\omega_{jst}$, the bias induced by endogeneity between $H$ and $\omega_{jst}$ is likely positive implying that the coefficient estimate will be biased upward ($\widehat{\beta_H} > \beta_H$).

In Section 6.1, I employ the productivity literature's techniques to control for the endogeneity of inputs in order to assess the relevance of this issue in my setting. In particular, I apply the Olley and Pakes (1996, henceforth OP) and the Levinsohn and Petrin (2003, henceforth LP) approaches. OP construct an explicit model for the firm's optimization problem in order to obtain their production function estimator. Essentially, the authors address the issue of endogeneity of inputs by using information about observed investment to proxy for unobserved productivity and by applying a control function estimator. Building on OP, LP suggest the use of intermediate input demand in place of investment demand as a proxy for unobserved productivity. Section A.1 contains a brief summary of the in-depth Eberhardt and Helmer (2010) discussion of these 'structural' estimators.

[16]Examples are shocks due to industry-level business cycles as well as changes in profit margins, industry concentration, and import competition.

## 4.3 Lemons' Bias

Another potential threat to identification is the fact that I do not observe labor quality. Of particular concern are workers who might separate from an HWF because their level of ability is low. The bias may work against the finding of a positive effect of knowledgeable workers.

In order to address this issue, I obtain a proxy for worker ability and I weight the number of workers in my OLS regression using the average ability to obtain effective labor input. Specifically, I weight the total number of workers $L_{jst}$ by firm $j$' average worker ability level $\overline{\theta}_{jt} = \frac{1}{L_{jt}} \sum_{i=1}^{L_{jt}} \theta_i$, to obtain effective labor input. $\theta_i$ is time-varying at firm level, given that the number and composition of workers change. To obtain $\theta_i$ I procure estimates of worker fixed effects from wage equations where both firm and worker effects can be identified. Section 4.4 describes this estimation in detail.

### 4.3.1 Using the number of downsizing firms as instrumental variable

To further guard against the possibility of lemons bias, I instrument for the number of knowledgeable workers in a non-HWF with the number of local good firms in the same industry that downsized in the previous period. I chose this particular instrument because descriptive evidence suggests a strong role for geographic and economic proximity in worker mobility. It may be argued that when a larger number of workers is laid off from good firms, lemons' bias is less likely to arise. One can think of two reasons why some good firms may downsize in a particular year. First, good firms may get a bad draw from the distribution of product-market conditions. Even though an inherent productivity advantage partly insulates the good firms from output increases, sufficiently large shocks will pierce this insulation and induce the good firm to layoff workers.[17] Alternatively, some good firms may also downsize in a particular year due to offshoring. Veneto businesses began to outsource their productive activities abroad in the mid-1990s. Several contemporaneous factors encouraged this phenomenon: the currency appreciation caused by Italy joining the Euro; an increase of competition at the international level; a global drop in transport costs and tariffs, and the burgeoning participation of Eastern European countries, Russia and China in the international consumption market. Veneto outsourcing of production abroad continued to grow over the past two decades.[18] If distance acts

[17] In the context of the simple model in Section 3, it would be interesting to adapt the analysis of plant closings in Hamermesh (1993) to permanent layoffs from continuing (good) enterprises. The conclusion of this paper discusses how one could further develop such a model.

[18] see Gianelle and Tattara, 2008, and Constantin, Giusti and Tattara, 2010, for analysis of the internationalization of Veneto firms.

as a barrier for job mobility then non-HWFs located in localities where good firms are downsizing in time $t$ will be more likely to hire a worker with HWF experience starting from $t+1$. In the presence of product demand shocks or offshoring, using the number of downsizing firms as an instrument is invalid if it cannot be excluded from the causal model of interest (Equation 3). The identifying assumption of my IV strategy is therefore that the number of downsizing good firms is correlated with the causal variable of interest, $H$, but uncorrelated with any other unobserved determinants of productivity.

## 4.4 Identification of Good Firms

I define good firms as high-wage-firms (HWFs). Following Abowd, Kramarz and Margolis (1999, henceforth, AKM), I specify a loglinear statistical model of wages as follows:

$$w_{ijt} = X'_{it}\beta + \theta_i + \psi_j + \upsilon_t + \varepsilon_{ijt} \tag{6}$$

where the dependent variable, the log of the average daily wage earned by worker $i$ in firm $j$ in year $t$, is expressed as a function of individual heterogeneity, firm heterogeneity, and measured time-varying characteristics. Firm and worker effects ($\psi_j$ and $\theta_i$) represent the earnings premium that a firm pays to each worker it employs, and the earnings premium that a worker receives in each firm she works for, respectively. The firm premium may reflect rent sharing, compensating differentials, or general heterogeneity across establishments in their compensation policies. The vector $X'_{it}$ includes tenure, tenure squared, age, age squared, a dummy variable for manager and white collar status, and interaction terms between gender and other individual characteristics). The assumptions for the statistical residual $\varepsilon_{ijt}$ are (a) $E[\varepsilon_{ijt}|i,t,x] = 0$, (b) $Var[\varepsilon_{ijt}|i,t,x] < \infty$ and (c) orthogonality to all other effects in the model.

The presence of labor mobility in matched worker-firm data sets enables the identification of worker and firm effects. The identification of each fixed effect in Equation (6) relies on the assumption that mobility is exogenous to the included regressors. Bias in the estimated firm effects arises when errors predict specific firm to firm transitions.[19]For my goal, this bias generates measurement error and works against finding any effect of recruiting knowledgeable workers on productivity.[20]The method in Abowd, Creecy and Kramarz (2002) identifies separate groups of workers and firms that are connected via

[19]If the idiosyncratic component of the errors only predicts separation but not specific transition, firm effects may not necessarily be biased.

[20]Card, Heining and Kline (2012) conduct a series of checks for patterns of endogenous mobility which could lead to systematic bias in AKM's additive worker and firm effects model. The authors find little evidence of such biases in German data.

labor mobility in the data. When a group of workers and firms is connected, the group contains all persons who ever worked for any firm within the group and all firms at which any of the persons were ever employed. In my fourteen-year sample, the largest group connected via mobility contains around 99% of the observations in the dataset. I run my estimation for the largest group, and define good firms as those whose estimated firm fixed effects falls within the top quintile of all estimated firm effects. See Section 5 for more details on the procedure.

# 5 Data and Descriptive Statistics

The data set is for Veneto, an administrative region in the Northeast of Italy with a population of around 5 million people (8 percent of the country's total). The region has undergone deep economic changes in the last few decades. Before World War II the economy was largely based upon farming and the region experienced large out-migration to Germany, Switzerland, the United States, Canada and Australia. The 1960s and 1970s, however, were characterized by intense economic development. Since the mid-1980s, the labor market in Veneto has been characterized by nearly full employment, a positive rate of job creation in manufacturing and positive migration flows (Tattara and Valentini, 2007). The dynamic regional economy, centred on manufacturing, features a large presence of flexible firms, frequently organized in districts with a level of industrial value added greatly exceeding the national average.[21]Manufacturing firms in Veneto specialize in metal-engineering, goldsmiths, plastics, furniture, garments, textiles, leather and shoes[22]. The manufacture of food and beverage, and wine and baked goods in particular, is also a prominent subsector.

My data set pools three sources of information: individual earnings records, firm balance sheets, and information on local labor systems[23].

The earnings records result from the Veneto Workers History (VWH) dataset. [24] The VWH has data on private sector personnel in the Veneto region over the period 1975-2001. Specifically, it contains register-based information for virtually any job lasting at least one day. A complete employment history has been reconstructed for each worker. For each

[21]The most famous industrial district is the eyewear district of Agordo, where Luxottica, the world's largest manufacturer of eyeglasses, has production plants.

[22]Benetton, Sisley, Geox, Diesel, and Replay are Venetian brands.

[23]The first two kinds of information, combined for the period 1995-2001, have been used in the study on rent-sharing, hold-up and wages by Card, Devicienti and Maida (2011).

[24]The VWH was assembled by Giuseppe Tattara and collaborators at University of Venice using administrative archives of the Istituto Nazionale per la Previdenza Sociale (INPS), the main public institute of Social Security in Italy. I am extremely grateful to Giuseppe Tattara for making the dataset available.

employee in the database, the VWH contains overall calendar-year earnings at each job, the number of days worked during the year, the relevant national contract and the level within that contract (i.e., a "job ladder" code), and the employee's age, gender, region (or country) of birth and total job tenure with each employer. For each firm, the VWH contains an industry categorization (assigned by five-digit code), start and closure dates (if applicable) of the firm and the firm's location[25].

Balance sheet records starting from 1995 were obtained from AIDA (Analisi Informatizzata delle Aziende), a database circulated by Bureau Van Dijk containing official balance sheet records of all incorporated non-financial Italian firms with annual revenues of at least 500,000 Euros. AIDA's balance sheets describe revenues, total wage bill, the book value of capital (broken into subgroups), value added, number of employees, value of materials and industry code for each firm.

I obtained Information on local labor markets (LLMs) from the National Institute of Statistics (ISTAT). The LLMs are territorial groupings of municipalities characterized by a certain degree of working-day commuting by the resident population. ISTAT conducted three studies on LLMs in 1981, 1991 and 2001. [26]Because my analysis uses examines the period 1987-2001 I utilize the LLM study from 1991. In this study the 518 municipalities or *comuni* in Veneto are divided into 51 LLMs.

I use firm identifiers to match job-year observations for workers aged 16-64 in the VWH with firm financial data in AIDA for the period 1995-2001. The match rate is fairly high: at least one observation in the VHW was found for over 95% of the employers in the AIDA sample, and around 50% of employees observed in the VWH between 1995 and 2001 can be matched to an AIDA firm. Most of the non-matches seem to be workers of small firms which are omitted from AIDA. In sum, I was able to match at least one employee for around 18,000 firms, or around 10% of the entire universe of employers contained in the VWH.[27]

From this set of potential matches I execute two exclusions to obtain my estimation sample for Equation (6). First, I remove all workers outside manufacturing. Further, I excluded job-year observations with remarkably high or low values for wages (I trim

[25] Additional information on the VWH is available in Battisti (2012), Tattara and Valentini (2011) and Card, Devicienti and Maida (2011).

[26] In the 2001 study 686 LLMs were recorded in Italy. The median number of working residents was 10,763. The average number of working residents was 30,576, with 1,251 and 1,321,564 recorded as the upper and lower bounds.

[27] Card, Devicienti and Maida (2011) show that the average firm size for the matched jobs sample (36.0 workers) is considerably larger than that for total employers in the VWH (7.0 workers). Mean daily wages for the matched observations are also greater, while the fractions of under 30 and female employees are lower.

observations outside the 1% - 99% range). As explained above, the method in Abowd, Creecy and Kramarz (2002) identifies separate groups of workers and firms in the data that are connected via labor mobility. I run the grouping algorithm separately using VHW data from 1987 to 2000 for firms that could be matched in AIDA. I then use the created group variable to choose the largest group as a sample for my fixed-effects estimation. The largest group contains 99.1% of the woker-year observations (2,567,040 observations combining 457,763 individuals with 5,937 firms). I identify HWFs those firms whose firm effects rank in the top 20% of the sample.[28] Descriptive statistics for HWFs in the sample are provided in Table A.1.The sample of non-HWFs used in the main firm-level analysis – equation (3) - is summarized in Table A.3. [29]

## 5.1 Characterization of Good Firms

For labor mobility to generate productivity benefits of agglomeration, a firm-specific advantage should be observed at good firms that could be the basis for knowledge transfer to other firms in the region. Therefore, once I have categorized firms into HWF and non-HWF groups, I estimate:

$$\ln O_{jst} = \beta_0 + \beta_1 HWF_{js} + \mu_s + v_t + e_{jst} \tag{7}$$

where the dummy $HWF$ takes the value of 1 if firm $j$ is classified as high-wage (a firm whose firm effect ranks in the top 20% of the sample) and $O_{jst}$ represents different firm-level outcomes, such as output per worker, value added per worker and tangible and intangible capital per worker.[30] Table 1 shows the results of estimating Equation (7) using the sample constructed for the fixed-effect estimation. In the Veneto manufacturing sector clear differences between HWFs and non-HWFs emerge in labor productivity (measured as

[28] In order to implement the approach in Abowd, Creecy and Kramarz (2002), I use the a2reg Stata routine developed by Ouazad (2007).

[29] In order to obtain this estimation sample I first remove HWF observations from the sample of worker-firm matches. From this non-HWF sample I remove (a) firms that close during the calendar year and (b) firm-year observations with remarkably high or low values (outside the 1% - 99% range) for several key firm-level variables, such as total value of production, number of employees, capital stock and value of materials. I then attempt to reduce the influence of false matches, particularly for larger firms, by implementing a strategy of Card, Devicienti and Maida (2011) to eliminate the "gross outliers", a minor number of matches (less than 1% of all employers) for which the absolute gap between the number of workers reported in a firm's AIDA balance sheet and the number found in the VWH is larger than 100.

[30] I could have defined good firms as the highly productive ones and detected them empirically using balance sheet data. There are two reasons why I chose not to pursue this strategy, and instead defined the good firms as HWFs. First, the availability of worker-level Social Security data allows the introduction of measured individual characteristics and worker effects impossible to capture with firm level data from balance sheets. Second, Social Security data are available for a longer period of time than the balance sheets, and therefore increase the precision of the categorization of firms into HWF and non-HWF groups.

output per worker, Column 1), value added per worker (Column 2) and firm size (Column 3). Table 1 also indicates differences in capital per worker (Column 4), including both tangible capital (Column 5) and, most remarkably, intangible fixed assets (Column 6). This evidence is important for establishing the potential for knowledge transfer in the region. Overall, these descriptive results point to an HWF advantage. In particular, since labor productivity is on average 16% higher in HWFs, and intangible capital per worker (intellectual property, accumulated research and development investments and goodwill) is 28% larger, we can also think of HWFs as high-productivity firms, or high-intangible-capital firms. Table A.2 illustrates that, in contrast to firm characteristics, workforce characteristics of HWFs and non-HWFs are not so different: the shares of white collar workers and managers are 2.8 and 0.4 percentage points higher, respectively in HWFs; the shares of female and workers older than 45 are 3.9 and 1.1 percentage points lower, respectively. No difference emerges in the share of workers younger than 30.[31]

## 5.2 The Extent of Labor Mobility

For labor mobility to be a mechanism for transfer of knowledge, we must observe some workers moving from HWFs to other firms. On average, between 1995 and 2001, 3.7% of non-HWFs in a given year employ workers with recent HWF experience. If we consider all non-HWFs that hire workers from HWFs (regardless of when the separation from the HWF took place), this percentage increases slightly to 4.1%. Overall, 835 workers switch from HWFs to non-HWFs during my sample period; 519 are blue collar workers, 284 are white collar workers, 20 are managers and 12 are apprentices. The vast majority of these workers - 812 of the 835 - have recent HWF experience.

It is important to observe that these numbers do not imply that in a typical year about 4 percent of Veneto firms are potentially affected by knowledge transfer. Recall that I only consider flows from firms in the top 20% of estimated firm fixed effects to firms in the bottom 80%. As a result, these numbers should be interpreted as implying that in a typical year about 4 percent of the firms in the bottom 80% of the distribution employ at least one worker with experience at a firm in the top 20%. There obviously exists significant labor mobility *within* the two groups that may also serve as a channel of knowledge transfer. To illustrate, one can intuitively imagine that a worker moving from a firm in the 1st percentile of the distribution to a firm in the 19th percentile may bring efficiency-enhancing knowledge to his or her new job[32], and the same can be imagined

[31] Since these specifications do not need require information collected from AIDA balance sheets, the sample period is not restricted to post-1995 observations.

[32] Despite potential lawsuits due to violations of non-compete covenants and trade secret law, one

for a worker moving from a firm in the 21st percentile to a firm in the 99th percentile. However in order to achieve clear identification I focus solely on flows *between* the two groups. In addition to the existence of labor mobility in my sample, I observe that the percentage of firms that employ workers with HWF experience varies with the threshold that I impose on the distribution. For instance, if I define HWFs as firms with fixed effects in the top 50% of the overall distribution, 8 percent of non-HWFs employ workers with HWF experience, compared with 4.1 percent if HWFs are defined by falling in the top 20% of the fixed-effects distribution.

As regards to individual characteristics of the movers in my sample, in all years movers from HWFs are significantly more likely to be young and male than non-HWFs workers without experience at good firms. In most years, these movers are also significantly more likely to be white-collar workers and managers. Table A.4 and A.5 give descriptive statistics in the most recent year (2001) for movers from good firms to non-HWFs and non-HWFs workers without experience at good firms.

In the next section I evaluate the extent to which non-HWFs benefit from hiring workers from HWFs by entering annual firm-level measures of the number of workers with experience at HWFs into a production function. Figure A.1 illustrates what I discussed above: $H = 0$ for the vast majority of the sample. The 95 percentile of the distribution is equal to 0. The mean number of H workers is 0.04, and the maximum is 3. [33]

# 6 Evidence on Worker Flows and Productivity

## 6.1 Main Production Function Estimates

In this section I evaluate the extent to which non-HWFs benefit from hiring workers from HWFs. Table 2 shows the results of estimating Equation (3) for the period 1995-2001. I cluster standard errors at the firm level. In Column 1, the coefficient on $H_{jst}$ indicates that a non-HWF's recruitment of a HWF worker is associated with an increase of 4.9 percent in the firm's productivity. [34] In Column 2, I augment the production function with $\widetilde{N}_{jst}$,the number of recently hired workers without experience from good firms. to address the issue

frequently observe top firms poaching employees from competitors in an effort to acquire some of their internal knowledge. This poaching is sometimes so intense that companies may cut deals to refrain from competing for employees. In December 2010, the U.S. Justice Department settled an antitrust suit with Lucasfilm over a "no solicitation" agreement with rival Pixar. In September of the same year, the Justice Department had settled another suit over similar agreements involving Adobe Systems, Apple, Google, Intel, Intuit and Pixar (*The New York Times*, January 2, 2011).

[33] The median number of employees at non-HWFs is 34.

[34] The overall production function has mild decreasing returns to scale, with a 1 percent increase in all inputs leading to a 0.91 percent increase in output.

of a potential correlation between recent hires and unobservables as per the discussion in Section 4.1. The coefficient on newly hired workers without HWF experience is highly significant and small (0.003). The difference in productivity effects associated with the each type of newly hired workers is highly significant. The productivity effect attributed to workers with HWF experience, therefore, does not appear to be associated with recently hired workers in general.

Column 3 and 4 of Table 2 present results that address the issue of transmission bias, i.e. the presence caused by unobservable shocks that 'transmit to' input choices, using the productivity literature's techniques. Column 3 reports results using the OP estimator.[35] In this specification $H_{jst}$ is treated as a freely variable input. The coefficient for $H_{jst}$ is positive (0.041) and significant. However, these estimates should be interpreted cautiously because I do not observe investment, and hence derived a proxy variable in $t$ as the difference between the reported book value of capital at time $t+1$ and its value in $t$. The way I constructed the proxy variable exacerbates the measurement error problems typically associated with the proxy variable approach. In addition, augmenting my specification with this proxy variable reduces my sample size substantially, as (a) 4628 out of 21539 firm-year observations are lost when I take the difference in reported book values and (b) the OP approach requires positive values for the proxy variable, eliminating an additional 8,546 firm-year observations.[36] Column 4 reports the results for LP estimator. [37] The coefficient for $H_{jst}$ is positive (0.028) and significant; it is lower than the OLS estimate, confirming the theoretical and empirical results on variable inputs discussed in LP. [38]Although the estimate of the coefficient for $H_{jst}$ in the OP and LP specification is smaller than the baseline estimate, none of the specifications is qualitatively inconsistent with the empirical finding that labor mobility works as a channel of knowledge transfer. Taken together, these estimates suggest that non-HWFs benefit from hiring workers with previous HWF experience by experiencing increased productivity.

Next, I address the questions of whether the knowledge embedded in workers is general enough to be applied in different industries, and whether the occupation of new

[35] I use the opreg Stata routine developed by Yasar, Raciborski, and Poi (2008).

[36] The OP estimation routine will truncate firms' non-positive proxy variable observations because the necessary monotonicity condition does not hold for these observations. See Section A.1 for more details.

[37] I use the *levpet* Stata routine developed by Levinsohn, Petrin, and Poi (2003).

[38] One expects the more variable inputs to be more strongly correlated with current values of productivity. As for effective labor $\overline{\theta}_{jt}L_{jst}$, the other variable input, the OLS estimate also exceeds the LP estimate. The results for capital are also consistent with LP. The authors show that if capital is not correlated with the current period's transmitted shock (but variable inputs are), or if capital has a much weaker correlation with the productivity shock than do the variable inputs, the OLS estimate on capital is likely to be biased downward.

hires affects the magnitude of the receiving firm's productivity benefits. Column 5 of Table 2 differentiates between workers with HWF experience moving within the same two-digit industry and workers moving between industries. The productivity gain from knowledgeable workers moving within industry is highly significant and positive (0.072). The gain from knowledgeable workers moving between industries is significant and positive but smaller (0.030). However, the difference in productivity changes associated with the two types of newly hired knowledgeable workers is not significant at conventional levels (p-value 0.128). Column 6 divides knowledgeable workers according to their occupation group in the current firms. The coefficients for both white-collar and blue-collar workers are positive (0.035 and 0.056) and significant, suggesting that the knowledge transferred through labor mobility is not exclusively patented or transferred by highly skilled workers. This is consistent with evidence on Danish manufacturing in Stoyanov and Zubanov (2012). The estimated coefficient for the dummy variable indicating managerial status (0.071) is large but not precisely estimated, likely due to the infrequency of managers to change jobs in my sample.[39]

### 6.1.1 Sensitivity analysis

The main empirical result in Section 6.1 is that labor mobility from HWFs to other firms in the region works as a mechanism for the transfer of efficiency-enhancing knowledge. This section investigates the robustness of these estimates to different specifications and explores potential alternative explanations of the estimated productivity effects. I begin by investigating the role of unobserved province shocks, localized industry shocks, unobserved firm heterogeneity and functional form assumptions. I conclude by evaluating the role of the selection of movers based on observable characteristics. Table A.6 shows results from a series of specification checks. As a basis for comparison, Column 1 shows the estimates from the baseline specification in Column 1 of Table 2. I begin by adding province-year fixed effects (Column 2) and by replacing industry-year fixed effects with province-year-industry fixed effects (Column 3). The goal of these two specifications is to purge the knowledge transfer effects of unobserved province-wide or province-by-industry shocks to productivity that might be correlated with the number of knowledgeable workers. The results are largely unchanged. Column 4 shows estimates using the within-transformation. These estimates should be interpreted cautiously because the within estimator is known from practical experience to perform poorly in the context of production functions (Eberhardt and Helmer, 2010). Indeed, estimates in Column 4

[39]The standard error for the coefficient on managerial status is 0.047; I observe only 20 managers moving from the top 20% group to the bottom 80%.

indicate severely decreasing returns to scale, likely due to measurement error in the input variables, whose influence is exacerbated by the variable transformation. The problem of using the within-transformation is the removal of considerable information from the data, since only variation over time is left to identify parameters. Setting this concern aside, the results show a positive and significant coefficient on $H$ (0.016) that is smaller than the baseline OLS coefficient, and the coefficients in other specifications reported in Table 2. Until now, I have presented results based on specifications where the intensity of potential knowledge transferred is measured by the number of H workers. In Column 5, I model this intensity as the share of workers with recent experience at good firms, which I denote with $h$. The coefficient is positive and significant: a one percentage point increase in $h$ is associated with a change in productivity of 0.9%. [40]

Considering the differences in observable characteristics documented in Section 5.2 between movers from HWFs and other workers at non-HWFs, I augment Equation (3) with the share of managers, females, white-collar workers, and differently aged workers at each firm. The results (reported in Column 2 of Table A.7) largely remained unchanged. Overall, the results in Table A.6 and A.7 are consistent with those discussed in Section 6.1.

## 6.2 The Role of Geographical Proximity

Having found strong evidence that labor turnover acts as a mechanism of knowledge transfer, I now analyze the potential role of geographic proximity in labor mobility, and consequently in the process of knowledge diffusion.

If labor mobility is a source of productivity advantages through agglomeration, geographic proximity should play a role in the process of knowledge diffusion. There exist at least two reasons why geographic proximity might be important in this context. First, distance may act as a barrier for workers' job mobility because of commuting costs or idiosyncratic preferences for location. [41]In January 2012, I visited several Veneto firms

[40]A potential problem with such specification arises, however, because there may be measurement error in $h$. In order to see this, rewrite Equation (3) as

$$\ln(\frac{Y_{jst}}{\bar{\theta} L_{jst}})=\beta_K \ln(K_{jst}) + \beta_M \ln(M_{jst}) + \beta_h h_{jst} + \mu_{st} + \lambda_t + \beta_0 + u_{jst}$$

Because $h = H/L$, a mechanical relantionship between $h$ and the dependent variable may arise at time $t$. To address this issue, I divide $H$ by the lagged number of employees reported in the social security data. The resulting coefficient, reported in Column 6, is positive (0.69) and significant at the 10% level.

[41]Descriptive statistics in Combes and Duranton (2006) show that labor flows in France are mostly local: about 75% of skilled workers remain in the same employment area when they switch firms. The

and interviewed employees about the history of their enterprises and their current operations. I also conducted phone interviews with officials of employers' associations and chambers of commerce. My anecdotal evidence supports the notion that distance acts as a barrier for job mobility.[42] Another reason geographical proximity may be an important determinant of job mobility is that the firm's informational cost of identifying the "right" employee are larger across localities than within them.[43]

### 6.3 Instrumental Variable Estimates

In Section 6.1, I addressed the issue of lemons' bias, i.e. the possibility that workers separate from an HWF because their level of ability is low, by weighting the number of workers using the average ability to obtain effective labor input. In this section I present estimates that use an instrumental variables (IV) strategy to address the same issue. I instrument for the number of knowledgeable workers using the number of good local firms in the same industry that downsized in the previous period. The identifying assumption of the IV strategy is that the number of downsizing good firms is correlated with the causal variable of interest, $H$, but uncorrelated with any other determinants of productivity. This exclusion restriction is violated if there are localized unobservable industry shocks that lead good firms to downsize and affect productivity at non-HWFs. A concern arises from the observation that the dependent variable in my econometric model is the *value* of output.[44] Unobserved shifts in local demand from HWFs to non-HWFs might lead to higher output prices, and hence higher productivity for non-HWFs. At the same time these shifts might lead HWFs to downsize and non-HWFs to hire HWF employees. I do not believe this to be a key issue in the context of my IV estimation; manufacturing firms in my sample generally produce goods traded outside the LLM.[45] Regardless, I

degree of geographical mobility implied by this figure is small, since the average French employment area is comparable to a circle of radius 23 kilometers. In Dal Bó, Finan and Rossi (2011), randomized job offers produce causal estimates of the effect of commuting distance on job acceptance rates. Distance appears to be a very strong (and negative) determinant of job acceptance: applicants are 33% less likely to accept a job offer if the municipality to which they are assigned is more than 80 kilometers away from their home municipality.

[42] In a phone interview, Federico Callegari of the Treviso Chambre of Commerce, reasoned out on the role of geographic proximity: "I think distance matters a lot for workers' job mobility. When losing their job, workers tend to look for another job with a commuting time of maximum 20-30 minutes. Why? Because they want to go home during the lunch break."

[43] A similar argument can be made for the informational costs for workers.

[44] The theoretically correct dependent variable in a productivity study is the *quantity* of output, but, due to data limitations, this study (and virtually all the empirical literature on productivity) uses price multiplied by quantity.

[45] Imagine the extreme case of a non-HWF that produces a nationally traded good in a perfectly competitive industry. Its output prices would not increase disproportionately if the LLM experienced an increased demand for its good.

address this potential issue by adding LLM-year fixed effects to control for unobserved local shocks.

Turning to the details of the instrument, I define downsizing of a firm as fulfilling two requirements. First, a downsizing firm must see an employment reduction equal or larger than 1 percent.[46] However, this division of good firms into downsizing and non-downsizing firms is less sensible for small firms. Accordingly, the second requirement is that the employment reduction must also equal or exceed three individuals. To summarize, the instrument is the lagged number of good firms in the same LLM and 3-digit industry of firm $j$ at time $t$ such that $(L_t^{good} - L_{t-1}^{good})/L_{t-1}^{good} \leq -0.01$ and $L_t^{good} - L_{t-1}^{good} \leq -3$.

Table 4 shows the results from the IV estimation of Equation (3). Standard errors are clustered at the level of the LLM. The F test of excluded instruments in Column 1 gives a statistic of 14.7. [47]The effect of $H$ on productivity is large: a unit increase in $H$ increases productivity by 31.2%. However, note that the standard errors are also large (14.5%). In column 2 I use the stricter definitions of downsizing firms such that a good firm is considered as downsizing if the decrease in the labor force is greater than three percent. [48] The estimated changes in productivity following a unit increase in $H$ is slightly smaller. Standard errors are still very large. In column 3 I increase the threshold to five percent. Results are very similar. Recall the OLS estimates: (a) impact on productivity of the recruitment of a knowledgeable worker is equal to 4.9 percent, and (b) impact of a knowledgeable worker moving within the same two-digit industry is unlikely to be larger than 7.2 percent. In principle, the IV estimates (that are likely to be driven by flows *within* industries, given the way the instrument is designed) are consistent with the idea that the OLS coefficient is biased downward. In practice however the IV standard errors are large and prevent me from drawing definitive conclusions.

Another tentative explanation for the magnitude of the IV results is that the effect of knowledgeable workers may be heterogeneous across firms. If there are indeed heterogeneous effects of $H$ on productivity, then consistent OLS measures the average effect of $H$ on productivity across all firms, while Two Stage Least Squares (TSLS) estimates the average effect in the subset of firms that are marginal in the recruitment decision, in the sense that they recruit knowledgeable workers if and only if there exists excess local supply. [49] If the effect of knowledgeable workers on productivity is larger for non-HWFs

[46] I also present results when this threshold is increased to 3 and 5 percent

[47] The coefficient of the number of downsizing firm in the first-stage regression if equal to 0.018 (standar error is 0.005). A one standard deviation increase in the intrumental variable is a associated with an increase in $H$ of 0.02.

[48] I still impose the additional condition that the decrease in the labor force is greater than or equal to three units. Both instrumental variables are summarized in Table 3

[49] See Imbens and Angrist (1994). For a recent application, see Eisensee and Stromberg, (2008)

that are marginal in the recruitment decision, the TSLS estimates will exceed those of consistent OLS.

# 7 To what extent can worker flows explain agglomeration advantages?

In this Section, I address the question of the extent to which worker flows can explain the productivity advantages of agglomeration. I relate my findings to the existing evidence on the productivity advantages of agglomeration, and the study by Greenstone, Hornbeck and Moretti (2010, henceforth, GHM) in particular. The authors find that incumbent plants in U.S. counties who successfully competed with other counties to attract a large manufacturing plants experienced significantly higher total factor productivity (TFP) after the plant's opening than did incumbent plants in counties that survived a long selection process but narrowly lost the competition. In order to evaluate to what extent worker flows might explain evidence on the productivity advantages of agglomeration, I simulate an event analogous to that studied by GHM but within my framework, and predict the change in local productivity that is due to labor mobility. The event is an increase in the number of good firms such that the change in local output is comparable to the output of the average large plant whose opening is considered by GHM[50].

A broad overview of my procedure is as follows. First, I estimate the effect on the number of $H$ workers *moving within industry* observed at firm $j$ of a change in the number of good local firms within the same industry as $j$. Denote the number of $H$ workers moving within industry with $H^{ind}$. Recall that for a worker to be counted as having recent HWF-experience in year $t$, the worker must be observed in a HWF for one or more of the years between $t-k$ and $t-1$. Then, if a worker is hired at time $t-g$, where $g \leq k$, she is counted as a knowledgeable worker from year $t-g$ until $t$.[51] This implies that $H^{ind}$ exhibits a certain degree of persistence and suggests estimation of a dynamic model for the number of workers observed at firm $j$.who have HWF experience in the same industry.

[50] The large plants in GHM generated bidding from local governments, almost certainly because there was a belief of important positive effects on the local economy. GHM observe that t he mean increase in TFP after the opening is (a) increasing over time and (b) larger if incumbent plants have the same industrial classification as the large plant. These two facts are consistent with the presence of intellectual externalities that are embodied in workers who move from firm to firm. I think of the establishments considered by GHM as "good" establishments, and in order to simulate their experiment I consider a change in the number of Venetian good firms such that the change in local output is comparable.

[51] It may be instructive to consider a practical example. Consider a worker who separates from a HWF in 1992 and joins non-HWF $j$ in 1995. Recall that in the baseline specification $k$ is set equal to 8. Provided that the worker remains in $j$, she will be counted as a knowledgeable worker for every year from 1995 to 2000.

In the second step of my simulation, I predict the change in $H^{ind}$ that each of the non-HWF in a LLM would experience if an output increase similar to the one considered by GHM were to occur, and I multiply the predicted change in $H^{ind}$ by $\widehat{\beta_H}^{ind}$,the estimated coefficient on $H^{ind}$ in my productivity regression. This product yields the predicted change in productivity for a given Veneto firm if its locality and industry were to experience an increase in output analogous to that considered by GHM.

In the final step of my simulation, I compare my back-of-the-envelope estimate from Step 2 of the predicted contribution of worker flows to productivity changes with GHM's estimate of the overall productivity effect, in order to have a sense of the extent to which worker flows can explain existing evidence on agglomeration advantages.

I will now discuss the issues related to the implementation of the first step, i.e. the estimation of the dynamic effect on $H$ of a change in the number of good firms in the same locality and industry.

Consider a model of the form

$$H^{ind}_{1jlst} = \widetilde{\alpha} H^{ind}_{jsl,t-1} + \widetilde{\beta} good\_firms_{ls(j)t} + \widetilde{\varepsilon}_{jlst} \tag{8}$$

$$\widetilde{\varepsilon}_{jlst} = \widetilde{\mu}_j + \widetilde{v}_{jlst} \tag{9}$$

$$E[\widetilde{\mu}_j] = E[\widetilde{v}_{jlst}] = E[\widetilde{\mu}_j \widetilde{v}_{jlst}] = 0 \tag{10}$$

where the dependent variable is the number of HWF-experienced workers observed at non-HWF $j$ in locality $l$ and industry $s$ at time $t$, and $good\_firms_{ls(j)t}$ is the number of good firms in the same locality *and same industry* of firm $j$. The disturbance term $\widetilde{\varepsilon}_{jlst}$ has two orthogonal components: the firm effect $\widetilde{\mu}_j$, and the idiosyncratic shock, $\widetilde{v}_{jlst}$. Using OLS to estimate Equation (8) is problematic because the correlation between $H^{ind}_{jsl,t-1}$ and the firm effect in the error term gives rise to "dynamic panel bias" (Nickell 1981). Application of the Within Groups estimator would draw the firm effects out of the error term, but dynamic panel bias would remain (Bond, 2002). Therefore I employ the first-difference transform, proposed by Arellano and Bond (1991). Applying this transform to Equation (8) yields:

$$\Delta H^{ind}_{1jlst} = \widetilde{\alpha} \Delta H^{ind}_{jsl,t-1} + \widetilde{\beta} \Delta good\_firms_{ls(j)t} + \Delta \widetilde{v}_{jlst} \tag{11}$$

The firm effects have now disappeared, but the lagged dependent variable is still potentially endogenous as the $H^{ind}_{jsl,t-1}$ in $\Delta H^{ind}_{jsl,t-1} = H^{ind}_{jsl,t-1} - H^{ind}_{jsl,t-2}$ is correlated with the $\widetilde{v}_{jls,t-1}$ in $\Delta \widetilde{v}_{jlst} = \widetilde{v}_{jls,t} - \widetilde{v}_{jls,t-1}$. However, longer lags of the regressors remain orthogonal to the error, and available for use as instruments. Natural candidate instruments for $\Delta H^{ind}_{jsl,t-1}$ are $H^{ind}_{jsl,t-2}$ and $\Delta H^{ind}_{jsl,t-2}$ . Both $H^{ind}_{jsl,t-2}$ and $\Delta H^{ind}_{jsl,t-2}$ are mathematically related to $\Delta H^{ind}_{jsl,t-1} = H^{ind}_{jsl,t} - H^{ind}_{jsl,t-1}$ but not to the error term $\Delta\widetilde{\upsilon}_{jlst} = \widetilde{\upsilon}_{jls,t} - \widetilde{\upsilon}_{jls,t-1}$ provided that the $\widetilde{\upsilon}_{jlst}$ are not serially correlated[52]. I use the classic Arellano-Bond Difference GMM estimator to estimate Equation (11) [53]. Table 5 gives the results of estimating Equation (11) for the period 1989-2001.[54] Column 1 shows a positive (0.008) and significant coefficient of the number of good local firms, and a positive (0.145) and significant coefficient for the lagged dependent variable. The p value of the Hansen test for overidentifying restrictions does not suggest misspecification. The Arellano-Bond test for serial correlation fails to indicate that the $\widetilde{\upsilon}_{jlst}$ are serially correlated [55]. Column 2 adds industry-year interaction terms to the baseline specification. The estimates for the coefficients on the number of good local firms, and the lagged dependent variable are very similar to the baseline estimates[56].

I can now move on to Step 2: predicting the changes in $H$, and hence in productivity, each of the non-HWF in Veneto would experience after an output increase similar to the one considered by GHM. As it turns out, the large manufacturing plants whose openings are studied by GHM are much larger than the median good firm in Veneto. [57] In order to observe a change in local output comparable to the average output increase caused by the opening of one large plant in GHM, a Veneto locality must experience an increase of 58 HWFs on average. Note in Equation (11) that a permanent unit change in $good\_firms_{ls(j)}$ yields on impact a change of $\widetilde{\beta}$, the next period it yields a change of $\widetilde{\alpha}\widetilde{\beta}$, and so on. The predicted change in $H$ that each non-HWF in Veneto would experience after 5 years, the time horizon considered in GHM, is then $\widehat{\Delta H}^{ind,5\ years} = 58\cdot(\widetilde{\beta}+\widetilde{\alpha}\widetilde{\beta}+\widehat{\widetilde{\alpha}}^{2}\widehat{\widetilde{\beta}}+\widehat{\widetilde{\alpha}}^{3}\widehat{\widetilde{\beta}}+\widehat{\widetilde{\alpha}}^{4}\widehat{\widetilde{\beta}}+\widehat{\widetilde{\alpha}}^{5}\widehat{\widetilde{\beta}})$. Therefore, in order to obtain the predicted change in

[52] Arellano and Bond (1991) develop a test for autocorrelation in the idiosyncratic disturbance term $\widetilde{\upsilon}_{jlst}$. I employ this test below.

[53] Kiviet (1995)'s approach is an alternate way to handle dynamic panel bias but does not address the potential endogeneity of other regressors. See Roodman (2006) for a detailed comparison of different approaches to estimate models which yield a reduced form with a lagged dependent variable.

[54] I include time dummies in order to remove universal time-related shocks from the errors.

[55] The Arellano-Bond test checks for serial correlation of order $l$ in levels by looking for correlation of order $l+1$ in differences

[56] The p-value of the Hansen test in Column 2 (0.770) is much larger than in Column 1 (0.246), probably because the addition of the interaction dummies causes an increase in the number of instruments. Roodman (2006) discusses how instrument proliferation can overfit endogenous variables and suggests that large values of the Hansen test p-value may be potential signs of trouble.

[57] This is due both to the fact that new entrants in GHM are significantly larger than the average new plant in the United States, and the fact that the Veneto region is characterized by the presence of small and medium-sized businesses, whose average size is smaller than the typical firm in United States. The median value of sales for HWFs in my sample is 6,787,000 of year-2000 euros. GHM show (Table 1, p. 555) that the mean output of large plants in their sample, five years after opening, equals 395,476,000 of 2000 Euros.

productivity for a given non-HWF if its locality and industry experience an output increase of the magnitude considered by GHM, I multiply $\widehat{\Delta H}^{ind,5\ years}$ by $\widehat{\beta_H}^{ind}$,the coefficient on knowledgeable in the baseline OLS productivity regression, estimated in Column 5 of Table 2. The predicted change in productivity attributable to worker flows five years the local output increase, the time horizon considered by GHM, is equal to $\widehat{\Delta TFP}^{ind,5\ years} = \widehat{\Delta H}^{ind,5\ years} \cdot 0.071 = 0.037$. In words,the average non-HWF experiences a 3.7% change in productivity.

The third step is to compare the magnitude of $\widehat{\Delta TFP}^{ind,5\ years}$, my back-of-the-envelope estimate of the predicted change in productivity due to worker flows with GHM's estimate of the overall productivity effect caused by a local output increase. The increase in productivity estimated by GHM five years after the opening for incumbent plants in the same two-digit industry equals 17 percent. Hence, my back of the envelope calculations suggest that worker flows explain 22 percent of the agglomeration advantages estimated by GHM.

Replacing $\widehat{\beta_H}^{ind,OLS}$ with $\widehat{\beta_H}^{ind,LP}$, the average effect of recruiting a knowledgeable worker with experience in the same industry estimated in the LP specification, the contribution of worker flows to the agglomeration advantages estimated by GHM is equal to 14 percent.[58] Overall, the back-of-the-envelope calculations in this section of the paper suggest that worker flows explain an economically relevant proportion of agglomeration advantages.

# 8 Conclusions and Future Directions

The evidence provided in this paper suggests that labor mobility from HWFs to other firms in Veneto manufacturing works as a mechanism for the transfer of efficiency-enhancing knowledge. I first showed that HWFs feature higher labor productivity, higher value added per worker and higher intangible capital per worker. This suggests that HWFs have a firm-specific advantage and hence, that there is a potential for knowledge transfer. I then showed that non-HWFs hiring workers with previous experience at HWFs benefit substantially in terms of increased productivity. Finally, I conducted back-of-the-envelope calculations suggesting that worker flows explain an economically significant portion of the productivity advantages of agglomeration estimated in a the well-known study by GHM.

There are several directions this work could take.

[58] The average effect of recruiting a knowledgeable worker with experience in the same industry estimated in the LP specification is $\widehat{\beta_H}^{ind,LP} = 0.044$ with standard error equal to 0.015 (not shown).

I first want to complement my Italian analysis with a similar exercise for the United States, using U.S. Census Bureau's Annual Survey of Manufactures and Longitudinal Employer-Household Dynamics data.

Moreover, I want to empirically analyze the extent to which workers' experience from good firms is rewarded in their new firms. If hiring non-HWFs do not fully pay for the value of workers to the firm, labor mobility is a source of a true knowledge externality. If instead the non-HWF fully pay for this value, there is no market failure. On a related note, I would like to further develop my model for the purpose of executing a comprehensive theoretical analysis of the labor market and spatial equilibrium in this economy. I plan to adopt a framework that extends the classic general equilibrium model of Rosen (1979) and Roback (1982) to allow for taste heterogeneity and commuting costs (Busso, Gregory and Kline, 2012), multiple industries, heterogeneity in firm productivity, downsizing at good firms and knowledge transfer.

Finally, I could employ additional instrumental variable approaches. For example I could instrument for labor mobility from HWF $a$ to non-HWF $b$ with a dummy variable that equals 1 if $a$ and $b$ rely on the services of the same temporary employment agency.[59].

These areas are all being actively pursued.

[59] Gianelle (2011) finds that temporary employment agencies significantly increased the integration and practicability of the inter-firm network of worker mobility in Veneto.

# References

[1] H. Abdel-Rahman and M. Fujita. Product variety, mashallian externalities, and city states. *Journal of Regional Science*, 30(2):165–183, 2006.

[2] J.M. Abowd, R.H. Creecy, and F. Kramarz. Computing person and firm effects using linked longitudinal employer-employee data. Technical report, Longitudinal Employer-Household Dynamics, Center for Economic Studies, US Census Bureau, 2002.

[3] J.M. Abowd, F. Kramarz, and D.N. Margolis. High wage workers and high wage firms. *Econometrica: Journal of the Econometric Society*, 67(2):251–333, 2003.

[4] D. Acemoglu. Training and innovation in an imperfect labour market. *The Review of Economic Studies*, 64(3):445–464, 1997.

[5] M. Arellano and S. Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297, 1991.

[6] R. Balsvik. Is labor mobility a channel for spillovers from multinationals? evidence from norwegian manufacturing. *The Review of Economics and Statistics*, 93(1):285–297, 2011.

[7] M. Battisti. *High wage workers and high wage peers*. PhD thesis, Simon Fraser University, 2011.

[8] N. Bloom, M. Schankerman, and J. Van Reenen. Identifying technology spillovers and product market rivalry. Technical report, National Bureau of Economic Research, 2007.

[9] S.R. Bond. Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, 1(2):141–162, 2002.

[10] S. Brusco. The emilian model: productive decentralisation and social integration. *Cambridge Journal of Economics*, 6(2):167–184, 1983.

[11] M. Busso, J. Gregory, and P.M. Kline. Assessing the incidence and efficiency of a prominent place based policy. Technical report, National Bureau of Economic Research, 2010.

[12] D. Card, F. Devicienti, and A. Maida. Rent-sharing, holdup, and wages: Evidence from matched panel data. Technical report, National Bureau of Economic Research, 2010.

[13] D. Card, J. Heining, and P. Kline. Workplace heterogeneity and the rise of german wage inequality. *Mimeograph UC Berkeley*, 2012.

[14] M.X. Chen and L. Alfaro. The global agglomeration of multinational firms. *The George Washington University, Institute for International Economic Policy Working Paper*, 2010.

[15] F. Cingano and F. Schivardi. Identifying the sources of local productivity growth. *Journal of the European Economic Association*, 2(4):720–744, 2004.

[16] P.P. Combes and G. Duranton. Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics*, 36(1):1–28, January 2006.

[17] F. Constantin, G. Giusti, and G. Tattara. Strategies of italian firms in romania: Evidence from selected case studies. *Transition Studies Review, Springer*, 16(4):829–847, February 2010.

[18] D.P. Cooper. Innovation and reciprocal externalities: information transmission via job mobility. *Journal of Economic Behavior and Organization*, 45(4):403–425, 2001.

[19] E. Dal Bó, F. Finan, and M.A. Rossi. Strengthening state capabilities: The role of financial incentives in the call to public service. *Mimeograph UC Berkeley*, 2011.

[20] K. Dasgupta. Learning and knowledge diffusion in a global economy. *Journal of International Economics*, 2011.

[21] G. Duranton and D. Puga. Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics*, 4:2063–2117, 2004.

[22] M. Eberhardt and C. Helmers. *Untested Assumptions and Data Slicing: A Critical Review of Firm-Level Production Function Estimators*. Department of Economics, University of Oxford, November 2010.

[23] T. Eisensee and D. Strömberg. News droughts, news floods, and us disaster relief. *The Quarterly Journal of Economics*, 122(2):693–728, 2007.

[24] G. Ellison, E.L. Glaeser, and W. Kerr. What causes industry agglomeration? evidence from coagglomeration patterns. Technical report, National Bureau of Economic Research, 2007.

[25] A. Fosfuri, M. Motta, and T. Rønde. Foreign direct investment and spillovers through workers' mobility. *Journal of International Economics*, 53(1):205–222, 2001.

[26] A. Fosfuri and T. Rønde. High-tech clusters, technology spillovers, and trade secret laws. *International Journal of Industrial Organization*, 22(1):45–65, 2004.

[27] M. Fujita and H. Ogawa. Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics*, 12(2):161–196, 1982.

[28] H. Gersbach and A. Schmutzler. Endogenous technological spillovers: causes and consequences. *Journal of Economics and Management Strategy*, 12(2):179–205, 2003.

[29] C. Gianelle. Exploring the complex structure of labour mobility networks: Evidence from veneto microdata. *Ca'Foscari University of Venice Working Paper*, 2011.

[30] C. Gianelle and G. Tattara. Manufacturing abroad while making profits at home: the veneto footwear and clothing industry. *Corporate Governance, Organization and the Firm: Co-operation and Outsourcing in the Global Economy*, page 206, 2009.

[31] E. Glaeser. Learning in cities. Technical report, National Bureau of Economic Research, 1997.

[32] E.L. Glaeser and J.D. Gottlieb. The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. Technical report, National Bureau of Economic Research, 2009.

[33] A.J. Glass and K. Saggi. Multinational firms and technology transfer. *The Scandinavian Journal of Economics*, 104(4):495–513, 2002.

[34] H. Görg and D. Greenaway. Much ado about nothing? do domestic firms really benefit from foreign direct investment? *The World Bank Research Observer*, 19(2):171–197, 2004.

[35] H. Görg and E. Strobl. Spillovers from foreign firms through worker mobility: An empirical investigation. *The Scandinavian Journal of Economics*, 107(4):693–709, 2005.

[36] M. Greenstone, R. Hornbeck, and E. Moretti. Identifying agglomeration spillovers: evidence from million dollar plants. Technical report, National Bureau of Economic Research, 2008.

[37] D.S. Hamermesh. *Labor Demand.* Princeton University Press, 41 William Street, Princeton, New Jersey 08540, 1993.

[38] R.W. Helsley. Knowledge and production in the cbd. *Journal of Urban Economics*, 28(3):391–403, 1990.

[39] R.W. Helsley and W.C. Strange. Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics*, 20(2):189–212, 1990.

[40] J.V. Henderson. Marshall's scale economies. *Journal of Urban Economics*, 53(1):1–28, 2003.

[41] G.W. Imbens and J.D. Angrist. Identification and estimation of local average treatment effects. *Econometrica: Journal of the Econometric Society*, pages 467–475, 1994.

[42] B. Jovanovic and Y. Nyarko. The transfer of human capital. *Journal of Economic Dynamics and Control*, 19(5):1033–1064, 1995.

[43] B. Jovanovic and R. Rob. The growth and diffusion of knowledge. *The Review of Economic Studies*, 56(4):569–582, 1989.

[44] U. Kaiser, H.C. Kongsted, and T. Rønde. Labor mobility and patenting activity. *University of Copenhagen, Department of Economics, Centre for Applied Microeconometrics (CAM) Working Paper*, 7, 2008.

[45] S. Kim. Labor specialization and the extent of the market. *Journal of Political Economy*, pages 692–705, 1989.

[46] J.F. Kiviet. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, 68(1):53–78, 1995.

[47] P. Kline. Place based policies, heterogeneity, and agglomeration. In *American Economic Review: Papers and Proceedings*, volume 100, pages 383–387, 2010.

[48] P. Kline and E. Moretti. Local economic development, agglomeration economies and the big push: 100 years of evidence from the tennessee valley authority. Technical report, UC Berkeley Working Paper, 2011.

[49] J. Levinsohn and A. Petrin. Estimating production functions using inputs to control for unobservables. Technical report, National Bureau of Economic Research Working Papers, 2000.

[50] R.E. Lucas Jr. On the size distribution of business firms. *The Bell Journal of Economics*, pages 508–523, 1978.

[51] M. Maliranta, P. Mohnen, and P. Rouvinen. Is inter-firm labor mobility a channel of knowledge spillovers? evidence from a linked employer–employee panel. *Industrial and Corporate Change*, 18(6):1161–1191, 2009.

[52] J.R. Markusen. Contracts, intellectual property rights, and multinational investment in developing countries. *Journal of International Economics*, 53(1):189–204, 2001.

[53] J.R. Markusen and N. Trofimenko. Teaching locals new tricks: foreign experts as a channel of knowledge transfers. *Journal of Development Economics*, 88(1):120–131, 2009.

[54] E. Moretti. Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review*, pages 656–690, 2004.

[55] E. Moretti. Local labor markets. *Handbook of Labor Economics*, 4:1237–1313, 2011.

[56] S. Nickell. Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426, 1981.

[57] G.S. Olley and A. Pakes. The dynamics of productivity in the telecommunications equipment industry. Technical report, National Bureau of Economic Research, 1992.

[58] A. Ouazad. Program for the estimation of two-way fixed effects. http://personal.lse.ac.uk/ouazad.

[59] H.G. Overman and D. Puga. Labor pooling as a source of agglomeration: An empirical investigation. In *Agglomeration Economics*, pages 133–150. University of Chicago Press, 2010.

[60] P. Parrotta and D. Pozzoli. The effect of learning by hiring on productivity. *The RAND Journal of Economics*, 43(1):167–185, 2012.

[61] A. Petrin, B.P. Poi, and J. Levinsohn. Production function estimation in stata using inputs to control for unobservables. *Stata Journal*, 4:113–123, 2004.

[62] MJ Piore, G. Becattini, M. Bellandi, and L. De Propis. Conceptualizing the dynamics of industrial districts. *The Handbook of Industrial Districts*, pages 259–268, 2009.

[63] M.J. Piore and C.F. Sabel. Italian small business development lessons for us industrial policy. *Massachusetts Institute of Technology Department of Economics Working Paper*, 1981.

[64] J.P. Poole. Knowledge transfers from multinational to domestic firms: evidence from worker mobility. *The Review of Economics and Statistics*, 2009.

[65] H. Rao and R. Drazin. Overcoming resource constraints on product innovation by recruiting talent from rivals: A study of the mutual fund industry, 1986-94. *Academy of Management Journal*, 45(3):491–507, 2002.

[66] J. Roback. Wages, rents, and the quality of life. *The Journal of Political Economy*, pages 1257–1278, 1982.

[67] D. Roodman. How to do xtabond2: An introduction to difference and system gmm in stata. *Center for Global Development Working Paper*, 2006.

[68] S. Rosen. Wage-based indexes of urban quality of life. *Current Issues in Urban Economics*, 1979.

[69] S.S. Rosenthal and W.C. Strange. Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, 85(2):377–393, 2003.

[70] S.S. Rosenthal and W.C. Strange. Evidence on the nature and sources of agglomeration economies. *Handbook of Regional and Urban Economics*, 4:2119–2171, 2004.

[71] J.J. Rotemberg and G. Saloner. Competition and human capital accumulation: a theory of interregional specialization and trade. *Regional Science and Urban Economics*, 30(4):373–404, 2000.

[72] A. Saxenian. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, 1996.

[73] J. Song, P. Almeida, and G. Wu. Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Science*, 49(4):351–365, 2003.

[74] A. Stoyanov and N. Zubanov. Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics*, 4(2):168–198, 2012.

[75] G. Tattara and M. Valentini. The cyclical behavior of job and worker flows working paper. *Department of Economics Ca'Foscari University of Venice*, 2007.

[76] G. Tattara and M. Valentini. Turnover and excess worker reallocation. the veneto labour market between 1982 and 1996. *Labour*, 24(4):474–500, 2010.

[77] C. Trigilia. Work and politics in the third italy's industrial districts. *Frank Pyke, Giacomo Becattini and Werner Sengenberger (eds) Industrial Districts and Inter-Firm Co-operation in Italy, Geneva: International Institute for Labor Studies*, pages 160–184, 1990.

[78] J. Whitford. The decline of a model? challenge and response in the italian industrial districts. *Economy and Society*, 30(1):38–65, 2001.

[79] M. Yasar, R. Raciborski, and B. Poi. Production function estimation in stata using the olley and pakes method. *Stata Journal*, 8(2):221, 2008 2008.

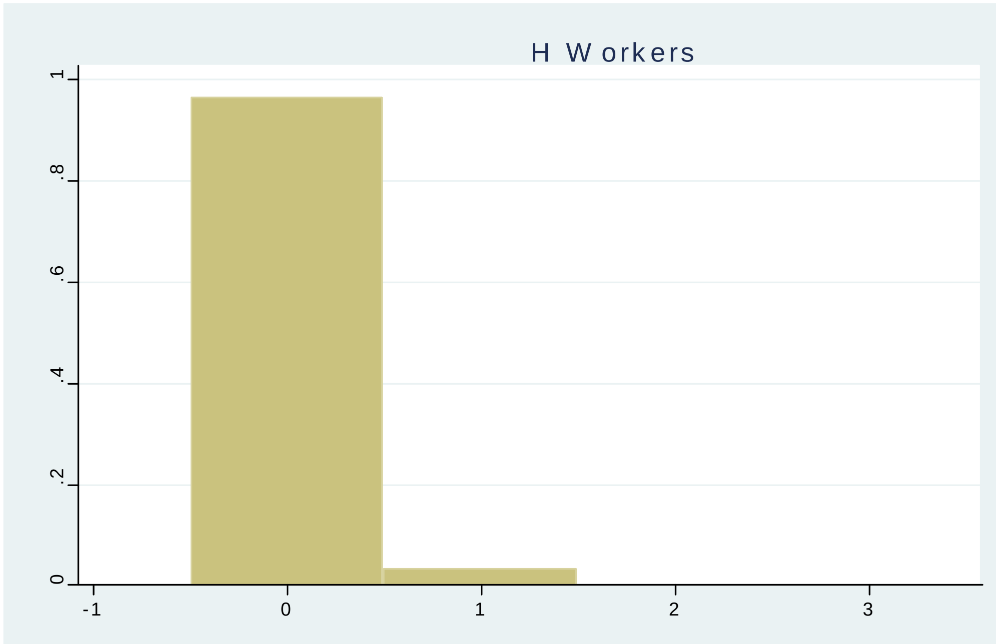


Table 1: Characteristics of HWFs, 1995-2001

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Y/L | VA/L | L | K/L | tang.K/L | intang.K/L |
| HWF | 0.153 | 0.126 | 0.036 | 0.090 | 0.048 | 0.276 |
| | (0.020) | (0.014) | (0.033) | (0.029) | (0.031) | (0.049) |
| Observations | 26041 | 26041 | 26041 | 26041 | 26041 | 26041 |

Dependent Variables are in Logs. All OLS regressions include year and industry dummies
Output, Value Added and Capital variables are in 1000's of 2000 real euros
Standard errors (in parentheses) clustered by firm
The dummy HWF takes value 1 if the firm is classified as high-wage

Table 2: H Workers and Productivity, 1995-2001

| | (1) Baseline | (2) Hires from non-HWF | (3) OP | (4) LP | (5) same/diff Ind | (6) by Occ |
|---|---|---|---|---|---|---|
| H workers | 0.049 | 0.040 | 0.041 | 0.028 | | |
| | (0.010) | (0.010) | (0.015) | (0.009) | | |
| log(capital) | 0.098 | 0.096 | 0.090 | 0.155 | 0.098 | 0.098 |
| | (0.004) | (0.004) | (0.017) | (0.009) | (0.004) | (0.004) |
| log(materials) | 0.574 | 0.572 | 0.575 | | 0.574 | 0.574 |
| | (0.007) | (0.007) | (0.008) | | (0.007) | (0.007) |
| log(employees) | 0.234 | 0.227 | 0.231 | 0.203 | 0.234 | 0.234 |
| | (0.007) | (0.007) | (0.009) | (0.005) | (0.007) | (0.007) |
| Hires from non-HWF | | 0.003 | | | | |
| | | (0.001) | | | | |
| H workers same Ind | | | | | 0.072 | |
| | | | | | (0.019) | |
| H workers diff Ind | | | | | 0.037 | |
| | | | | | (0.012) | |
| H managers | | | | | | 0.071 |
| | | | | | | (0.047) |
| H white collars | | | | | | 0.035 |
| | | | | | | (0.016) |
| H blue collars | | | | | | 0.056 |
| | | | | | | (0.015) |
| Observations | 21539 | 21539 | 8365 | 21539 | 21539 | 21539 |
| Adj. R-squared | 0.928 | 0.928 | | | 0.928 | 0.928 |

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm

Column 1 reports estimates from the baseline specification

Column 2 adds to Column 1 the number of newly hired workers from non-HWFs

Column 3 implements the procedure in Olley and Pakes (1996)

Column 4 implements the procedure in Levinsohn and Petrin (2003)

Column 5 differentiates between workers moving within the same industry and between industries

Column 6 divides H workers according to their occupation in the receiving firm

Table 3: Instrumental Variables

| **Variable** | **Mean** | **(Std. Dev.)** | **Min.** | **Max.** | **N** |
|---|---|---|---|---|---|
| lag1(downsizing HWFs, 3 percent drop in L) | 0.242 | (0.817) | 0 | 7 | 21475 |
| lag1(downsizing HWFs, 5 percent drop in L) | 0.224 | (0.746) | 0 | 6 | 21475 |

Table 4: H Workers and Productivity, IV Estimates 1995-2001

| | (1) (Log(Output) downsizing: 1 percent drop | (2) Log(Output) 3 percent drop | (3) Log(Output) 5 percent drop |
|---|---|---|---|
| H workers | 0.312 | 0.307 | 0.328 |
| | (0.145) | (0.147) | (0.163) |
| log(capital) | 0.098 | 0.098 | 0.098 |
| | (0.005) | (0.005) | (0.005) |
| log(materials) | 0.575 | 0.575 | 0.575 |
| | (0.009) | (0.009) | (0.009) |
| log(employees) | 0.227 | 0.228 | 0.227 |
| | (0.009) | (0.009) | (0.009) |
| Observations | 21475 | 21475 | 21475 |
| Adj.Rsquared | 0.916 | 0.916 | 0.916 |
| Fstat, instrum., 1st stage | 14.66 | 12.62 | 11.05 |

Dependent variable: Log(Output)

Standard errors (in parentheses) clustered by LLM (47)

Regressions include industry-year interaction dummies and LLM-year interaction dummies

Column 1 reports IV estimates using the lagged number of downsizing local good firms in the same 3-digit industry

In Column 2 a good firm is considered as downsizing if the drop in L is larger or equal than 3 percent

In Column 3 a good firm is considered as downsizing if the drop in L is larger or equal than 5 percent

Table 5: Number of local HWFs in same Industry and H Workers moving within Industry, 1989-2001

| | (1) Baseline | (2) Industry - Year FE |
|---|---|---|
| lag(H workers moving within industry) | 0.145 | 0.148 |
| | (0.048) | (0.048) |
| Local HWFs in same Ind | 0.008 | 0.007 |
| | (0.003) | (0.003) |
| Observations | 35723 | 35723 |
| AR(1)z | -7.481 | -7.480 |
| AR(2)z | 0.248 | 0.308 |
| HansPv | 0.246 | 0.770 |

Dependent variable: H Workers moving within Industry

Column 1 reports the baseline Difference GMM results. It includes year dummies

Column 2 includes industry-year interaction dummies

The variable Local HWFs in same Ind is treated as endogenous

Lags 2 to 3 of the variables are used as instruments

# A Appendix

## A.1 Structural Estimators of production functions

Several solutions for the endogeneity of input choices with regard to unobserved productivity have been proposed in the literature What follows is a brief summary of the in-depth discussion of 'structural' estimators in Eberhardt and Helmer (2010). OP address the issue of endogeneity of inputs by using information about observed investment to proxy for unobserved productivity and by applying a control function estimator. They assume that $k_{jt}$ and $\omega_{jt}$ are firm-specific state variables in the firm's dynamic programming problem. The Bellman equation is

$$V_{jt}(k_{jt,}\omega_{jt}) = \max\{\pi_j(k_{jt,}\omega_{jt}) - c_j(i_{jt}) + \theta E[V_{t+1}(k_{jt+1,}\omega_{jt+1})|k_{jt,}\omega_{jt}, i_{jt}]\}$$

where $k_{jt+1} = (1-\delta)k_{jt} + i_{jt}$ is the law of motion for capital accumulation. Investment

is chosen at time $t$ and adds to the capital stock at time $t+1$.The solution gives an investment policy function that depends on capital and productivity $i_{jt}(k_{jt,}\omega_{jt})$. Labor is not included in the investment equation because it is assumed to be a 'non-dynamic' input: it can be adjusted after realization of $\omega_{jt}$ within the same period. A key assumption is that investment is strictly increasing in both capital stock and productivity. In addition, $\omega_{jt}$ is assumed to be the only unobservable driving the investment choice. Finally, when deciding upon investment in period $t+1$ any realizations of $\omega_{jt}$ prior to time $t$ are not incorporated in the investment function because productivity evolves by assumption following an 'exogenous first-order Markov process': a firm builds expectations about its productivity at time $t+1$ exclusively based on its productivity levels realized at time $t$. Therefore one can assume most generally that productivity evolves according to $\omega_{jt} = g(\omega_{jt-1}) + \xi_{jt}$, where $\xi_{jt}$ is the random 'productivity shock'. Provided the investment function is continuous in $k_{jt,}$ and $\omega_{jt}$, and provided investment is positive, the investment equation can be inverted to yield $\omega_{jt} = f_t(i_{jt}, k_{jt})$. The OP estimator is implemented in two stages: first, by estimating

$$y_{jt} = \beta_l l_{jt} + \phi_{jt}(i_{jt}; k_{jt}) + \epsilon_{jt}$$

where

$$\phi_{jt}(i_{jt}; k_{jt}) = \beta_o + \beta_K k_{jt} + f_t(i_{jt}, k_{jt}) \tag{12}$$

OP propose estimation based on a third-order polynomial series expansion. In the second step, OP employ these estimates to run a regression of $y_{jt} - \widehat{\beta}_l l_{jt}$ on $\widehat{\phi}_{jt}(\cdot)$ and $k_{jt}$, which yields an unbiased $\widehat{\beta_K}$ . From the assumption of a Markov process for productivity and equation (12) one can realize that

$$E[\omega_{jt}|\omega_{jt-1}] = g(\phi_{jt-1}(i_{jt-1}; k_{jt-1}) - \beta_o - \beta_K k_{jt-1}) + \xi_{jt}$$

This allows one to write

$$y_{jt} - \widehat{\beta}_l l_{jt} = \beta_K k_{jt} + g(\widehat{\phi}_{jt-1}(i_{jt-1}; k_{jt-1}) - \beta_o - \beta_K k_{jt-1}) + \xi_{jt} + \epsilon_{jt} \tag{13}$$

Given that $\beta_K$ enters the equation twice and in combination with other parameters, equation (13) is estimated using non-linear least squares (NLLS).

The OP model can be extended to include firm exit, in which case an extra step is added between the two described above, where a probit regression is fitted on a nonlinear function of $i_{jt}$, $k_{jt}$ using the same argument of proxied productivity as in the first step. The predictions from this intermediate step are then added in the $g()$ function in the above second step.

Building on OP, LP suggested the use of intermediate input demand instead of investment demand as a proxy for productivity $\omega_{jt}$. This means that the decision on intermediate input is made at time $t$ once $\omega_{jt}$ is observed by the firm. The same applies to labor input choices, which in turn means that labor and intermediate inputs are chosen at the same time, and labor preserves its assumed non-dynamic/flexible nature. In the LP approach, intermediate inputs (electricity, material inputs) are modelled as a function of $\omega_{jt}$ and $k_{jt}$ similar to the use of investment in the OP procedure. See Eberhardt and Helmer (2010) for further details.

## A.2 Additional Tables

Table A.1: HWFs, Descriptive Statistics (1187 Individual Firms in 1987-2001)

| Variable | Mean | (Std. Dev.) | Min. | Max. | N |
|---|---|---|---|---|---|
| Output in 2000 (1000's of euros) | 18592.788 | (49759.703) | 544 | 861456 | 917 |
| Capital in 2000 (1000's of euros) | 4864.385 | (15280.61) | 1 | 170791 | 917 |
| Materials in 2000 (1000's of euros) | 10078.731 | (27125.103) | 0 | 412207 | 917 |
| Value added in 2000 (1000's real euros) | 4546.236 | (15782.815) | -1654 | 319641 | 917 |
| Tangigle capital in 2000 (1000's real euros) | 4235.313 | (13706.961) | 0 | 169724 | 917 |
| Intangible capital in 2000 (1000's real euros) | 629.072 | (4771.577) | 0 | 131417 | 917 |
| firm age (years) in 2000 | 17.603 | (13.092) | 1 | 117 | 917 |
| employees | 76.326 | (199.323) | 1 | 6188 | 6602 |
| apprentices | 0.663 | (2.676) | 0 | 138 | 10334 |
| blue collars | 38.574 | (101.555) | 0 | 3915 | 10334 |
| white collars | 16.615 | (61.029) | 0 | 1534 | 10334 |
| managers | 1.533 | (9.515) | 0 | 408 | 10334 |
| female employees | 13.459 | (62.925) | 0 | 2692 | 10334 |
| employees age$<$ 30 | 17.786 | (47.704) | 0 | 1616 | 10334 |
| employees age$>$ 45 | 14.365 | (39.843) | 0 | 795 | 10334 |

Table A.2: Characteristics of HWFs Workforce, 1987-2001

| | (1) share white coll. | (2) share manager | (3) female share | (4) share age<30 | (5) share age>45 |
|---|---|---|---|---|---|
| HWF | 0.028 | 0.004 | -0.039 | 0.004 | -0.011 |
| | (0.005) | (0.001) | (0.005) | (0.005) | (0.004) |
| Observations | 58102 | 58102 | 58102 | 58102 | 58102 |

All OLS regressions include year and industry dummies

Standard errors (in parentheses) clustered by firm

The dummy HWF takes value 1 if the firm is classified as high-wage

Table A.3: non-HWFs, Main Estimation Sample (4397 Individual Firms in 1995-2001)

| **Variable** | **Mean** | **(Std. Dev.)** | **Min.** | **Max.** | **N** |
|---|---|---|---|---|---|
| Output in 2000 (1000's of euros) | 8919.969 | (10406.331) | 1102 | 85372 | 3475 |
| Capital in 2000 (1000's of euros) | 2115.44 | (2914.85) | 61 | 23195 | 3475 |
| Materials in 2000 (1000's of euros) | 4608.517 | (6386.013) | 73 | 50939 | 3475 |
| Value added in 2000 (1000's real euros) | 2209.367 | (2553.203) | -2117 | 36787 | 3475 |
| Tangigle capital in 2000 (1000's real euros) | 1943.112 | (2747.577) | 20 | 22844 | 3475 |
| Intangible capital in 2000 (1000's real euros) | 172.328 | (503.237) | 0 | 10755 | 3475 |
| firm age (years) in 2000 | 18.552 | (11.089) | 0 | 99 | 3475 |
| employees | 50.556 | (51.332) | 2 | 450 | 21539 |
| apprentices | 1.01 | (1.978) | 0 | 47 | 21539 |
| blue collars | 31.348 | (32.665) | 0 | 365 | 21539 |
| white collars | 10.195 | (13.176) | 0 | 253 | 21539 |
| managers | 0.735 | (2.005) | 0 | 54 | 21539 |
| female employees | 13.406 | (19.859) | 0 | 309 | 21539 |
| employees age< 30 | 14.498 | (14.604) | 0 | 201 | 21539 |
| employees age> 45 | 9.572 | (13.679) | 0 | 199 | 21539 |
| H workers | 0.038 | (0.211) | 0 | 3 | 21539 |
| N workers | 3.877 | (7.595) | 0 | 212 | 21539 |
| H workers same Ind | 0.012 | (0.119) | 0 | 3 | 21539 |
| H workers diff Ind | 0.026 | (0.17) | 0 | 3 | 21539 |
| H managers | 0.002 | (0.041) | 0 | 1 | 21539 |
| H white collars | 0.013 | (0.123) | 0 | 3 | 21539 |
| H blue collars | 0.022 | (0.155) | 0 | 3 | 21539 |

Table A.4: H Workers in 2001

| **Variable** | **Mean** | **(Std. Dev.)** | **Min.** | **Max.** | **N** |
|---|---|---|---|---|---|
| age | 33.813 | (8.481) | 18 | 62 | 407 |
| female | 0.251 | (0.434) | 0 | 1 | 407 |
| blue collar | 0.548 | (0.498) | 0 | 1 | 407 |
| white collar | 0.388 | (0.488) | 0 | 1 | 407 |
| manager | 0.049 | (0.216) | 0 | 1 | 407 |

Table A.5: Workers without HWF experience in 2001

| **Variable** | **Mean** | **(Std. Dev.)** | **Min.** | **Max.** | **N** |
|---|---|---|---|---|---|
| age | 37.08 | (9.538) | 16 | 65 | 192588 |
| female | 0.32 | (0.467) | 0 | 1 | 192588 |
| blue collar | 0.71 | (0.454) | 0 | 1 | 192352 |
| white collar | 0.242 | (0.428) | 0 | 1 | 192352 |
| manager | 0.023 | (0.15) | 0 | 1 | 192352 |

Table A.6: Workers with HWF experience and Productivity, Robustness to Different Specifications

| | (1) Baseline | (2) Prov-Year FE | (3) Prov-Ind-Year FE | (4) Within | (5) Share | (6) H/lagL |
|---|---|---|---|---|---|---|
| H workers | 0.049 | 0.048 | 0.045 | 0.016 | | |
| | (0.010) | (0.010) | (0.010) | (0.006) | | |
| log(capital) | 0.098 | 0.098 | 0.101 | 0.064 | 0.098 | 0.100 |
| | (0.004) | (0.004) | (0.004) | (0.005) | (0.004) | (0.005) |
| log(materials) | 0.574 | 0.575 | 0.571 | 0.598 | 0.575 | 0.574 |
| | (0.007) | (0.007) | (0.007) | (0.012) | (0.007) | (0.008) |
| log(employees) | 0.234 | 0.234 | 0.236 | 0.062 | 0.235 | 0.244 |
| | (0.007) | (0.007) | (0.007) | (0.004) | (0.007) | (0.009) |
| share of H workers | | | | | 0.885 | |
| | | | | | (0.184) | |
| H/lagL | | | | | | 0.685 |
| | | | | | | (0.377) |
| Observations | 21539 | 21539 | 21539 | 21539 | 21539 | 16249 |
| Rsquared | 0.929 | 0.930 | 0.930 | 0.989 | 0.929 | 0.935 |

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm

All regressions include year dummies

Column 1 reports estimates from the baseline specification

Column 2 adds province-year interaction dummies

Column 3 replaces industry-year interaction dummies with province-industry-year interaction dummies

Column 4 reports within estimates

Column 5 replaces the number of H workers with the share of H workers

Column 6 uses the ratio of H workers over lagged number of employees

Table A.7: H Workers and Productivity, Robustness

| | (1) Baseline b/se | (2) Obs. Characteristics b/se |
|---|---|---|
| log(capital) | 0.098 | 0.095 |
| | (0.004) | (0.004) |
| log(materials) | 0.574 | 0.564 |
| | (0.007) | (0.007) |
| H workers | 0.049 | 0.043 |
| | (0.010) | (0.010) |
| log(employees) | 0.234 | 0.240 |
| | (0.007) | (0.007) |
| Observations | 21539 | 21539 |
| Adj. R-squared | 0.928 | 0.929 |

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm
Regressions include industry-year dummies
Column 1 reports estimates from the baseline specification
Column 2 adds share of managers, females, white collars, and differently aged workers